

Thomas Mann und Topic Maps. Potentiale neuer Informationstechnologien

Ingrid Schmidt, Carolin Müller

Herbst 1929: Der Knaur-Verlag will eine preiswerte Volksausgabe der Buddenbrooks in seiner Reihe der klassischen Werke der Weltliteratur herausgeben und sucht dafür um eine einmalige Sonderlizenz beim S. Fischer Verlag in Frankfurt am Main nach. Dies führte zu Meinungsverschiedenheiten zwischen Thomas Mann und seinem Verleger Samuel Fischer. Fischer war grundsätzlich gegen die Idee der Volksausgabe; Thomas Mann befürwortete sie und führte als Argument die „geänderte Zeit“ ins Feld. „Man müsse ihr Rechnung tragen, und die Volksausgabe eines Buchs, das seine Lebensfähigkeit ein Menschenalter lang bewährt und bewiesen habe, sei, wenn auch einem Verlagswesen wie S. Fischer ein Novum, dennoch beileibe nichts Ehrenrühriges, kein Ramschgeschäft, das den Buchhandel in seinen Festen erschüttern müsse, sondern Forderung der Zeit.“[3] Die Verhandlungen mit Knaur scheiterten; der Fischer Verlag brachte schließlich selbst eine Volksausgabe der Buddenbrooks heraus. Diese wurde ein Riesenerfolg.

Herbst 2001: Bei S. Fischer erscheinen die ersten drei Bände der *Großen kommentierten Frankfurter Ausgabe* der Werke Thomas Manns als Buchausgabe und CD-ROM. Im Vorfeld bewarben sich mehrere Softwarehersteller bei S. Fischer um die Lizenz für eine elektronische Version. Fischer entschied sich jedoch dafür, sie in eigener Regie herzustellen. [6] Diese Entscheidung leitete gleichzeitig einen Strategiewechsel im Verlag bezüglich der neuen Medien ein: Die Inhalte selber werden zur Sache des Verlags, unabhängig vom ihrem Publikationsmedium.

War 1929 die Volksausgabe eine Neuerung im Verlagsbereich, deren Möglichkeiten es zu nutzen galt, so sind es heute die elektronischen Medien, die den Verlagen neue Potentiale eröffnen. Mit der Veränderung der Medienlandschaft entstehen neue Publikationsmöglichkeiten neben – und nicht statt – der alten. Das Buch hat Gesellschaft bekommen: CD-ROMs, Internet, E-Books. Mehr denn je geht es heute um die Inhalte als solche. Sie werden nicht mehr ausschließlich aufs Papier gebracht; sie werden vielmehr zu entmaterialisierten Datenströmen, die idealerweise unabhängig von ihrem aktuel-

len Publikationsmedium sind und – wie immer auch die „Forderung der Zeit“ sein mag – in den unterschiedlichsten Medien manifest werden können.

Solche Möglichkeiten entstehen allerdings nicht von selbst, man muss sie schaffen. Dafür sind tragfähige Konzepte notwendig, die für die jeweiligen Inhalte maßgeschneidert sind, aber auch die praktische Umsetzbarkeit und die Kosten nicht aus den Augen verlieren; Konzepte, die zudem Standards mit Visionen verbinden. Standards spielen immer dann eine wichtige Rolle, wenn es darum geht, bei langfristig anzulegenden Planungen auf möglichst sicherem Boden zu stehen. Im Publikationsbereich sind SGML, XML [2] und alle ihre Verwandten hier an erster Stelle zu nennen. Visionen sind vonnöten, um heute schon zukunftsweisende Wege beschreiten zu können. Semantische Netze oder Topic Maps [5], ihre XML- bzw. SGML-Version, werden eine zunehmend wichtige Rolle spielen, wenn es darum geht, die Inhalte großer Datenmengen gezielt recherchierbar zu machen. Ein Sachverhalt, der heute als „knowledge management“, „wissensbasiertes Publizieren“ u.ä. Schlagzeilen macht.

Thomas Mann bei S. Fischer

Thomas Mann ist einer der wichtigsten deutschen Autoren der ersten Jahrhunderthälfte. In seinem umfangreichen Werk spiegeln sich viele große Themen seiner Zeit; es gehört zum festen Kanon der deutschen Literatur. Für den S. Fischer Verlag ist Thomas Mann bis heute einer der umsatzstärksten Autoren. Der Verlag hält bis 2025 die exklusiven Publikationsrechte an seinem Werk. Zwischen 2001 und 2015 wird nun bei Fischer in jährlichem Rhythmus eine 58-bändige Referenzausgabe der Werke, Briefe und Tagebücher Thomas Manns erscheinen. Diese *Große kommentierte Frankfurter Ausgabe* (GKFA) wird von über 15 international renommierten Wissenschaftlern herausgegeben und in etwa 20 Bänden umfassend kommentiert.

Die neue Referenzausgabe

Die Entscheidung zu dieser neuen Referenzausgabe – die über Jahrzehnte hinweg einen einzigartigen Zugang zu den Werken Thomas Manns bieten soll – ging im Verlag mit einer Reihe weiterer Überlegungen einher. Bei einem auf so lange Zeit hin angelegten Projekt lag es auf der Hand, die Ausgabe nicht nur als Buch, sondern auch elektronisch zu publizieren. Nach Sichtung der am Markt schon vorhandenen elektronischen literarischen Editionen war klar, dass diese Produkte nicht der Vorstellung einer elektronischen Version der GKFA entsprachen. Nicht klar war allerdings, wie genau eine elektron-

ische Ausgabe der GKFA aussehen sollte; auf keinen Fall war ein elektronisches Abbild des Buches angestrebt. Es sollte sichergestellt sein, dass die elektronische Ausgabe den hohen Qualitätsanforderungen der Buchausgabe entspricht. Qualitativ hochwertig hieß dabei neben hochwertigen Inhalten auch deren mediengerechte Umsetzung bei der Publikation. Darüber hinaus war gefordert, im Zuge dieser großen Ausgabe die Inhalte so aufzubereiten, dass sie langfristig nutzbare Ressourcen werden. Bei einem anderen großen Projekt im Fischer Verlag – dem Fischer Weltalmanach – hatte sich seit einigen Jahren SGML/XML für die langfristige Datenhaltung bewährt. [1] Dies war ein weiterer Grund dafür, auch für Thomas Mann auf diese Standards zu setzen. Daraus resultierten als zentrale Projektvorgaben, (1) dass die GKFA als Buchausgabe und als elektronische Ausgabe erscheinen wird, und (2) dass die Texte der GKFA SGML/XML-strukturiert vorliegen sollen.

Mit der ersten Projektvorgabe war nicht angestrebt, sich an das gängige Publikationsmodell für literarische Editionen anzulehnen. Herkömmlich entsteht die elektronische Version in der Regel auf der Datenbasis der gedruckten Ausgabe (Abbildung 1). Diese Datenbasis, die zunächst die Herstellung der Buchausgabe durchlaufen hat, transportiert folgerichtig in erster Linie typographische und layoutorientierte Informationen.

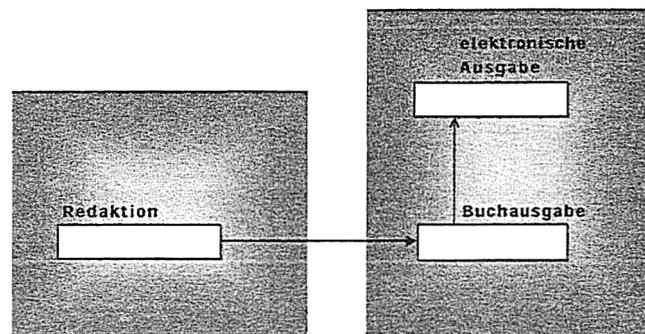


Abbildung 1: traditionelle Herstellung einer elektronischen literarischen Edition

Eine elektronische Version kann sich unter solchen Voraussetzungen nur mit erheblichem Zusatzaufwand von der Buchausgabe lösen. Die Folge ist ein elektronisches Buchimitat, das gerade wegen dieser Imitation erheblich an Qualität einbüßt: Es entsteht eine Art Buch, welches weder das klassische Medium Papier mit all seinen sinnlichen Eigenschaften benutzt, noch die Eigenschaften des elektronischen Mediums ausschöpft. Einzig über Volltextsuche und einige festgelegte semantische Suchmöglichkeiten wird eine Annäherung an die Möglichkeiten der neuen Medien erreicht. Das für die GKFA angestrebte Modell hingegen zielt gerade darauf ab, die spezifischen Eigenschaften – den Werkzeugcharakter – elektronischer Medien nutzbar zu ma-

chen. Gleichzeitig sollte sichergestellt sein, dass es sich um *eine* Referenzausgabe handelt, die in *zwei* Medien realisiert ist. Dabei ist das Buch die schöne Leseausgabe, die elektronische Version das Arbeitswerkzeug für den Literaturwissenschaftler.

Das verlegerische Herzstück

Die zweite Projektvorgabe zielte darauf ab, den Workflow des herkömmlichen Modells dahingehend zu verändern, dass die verschiedenen Publikationsmedien aus einer Datenbasis parallel bedient werden können. Die Datenbasis wird dabei in einem Informationspool abgespeichert. Dieser besteht (1) aus einer Textbasis, die ihrerseits SGML/XML-Instanzen, Textdaten in anderen Formaten, Grafiken, Musik etc. umfasst, (2) aus den zu den SGML/XML-Instanzen gehörigen DTDs sowie (3) aus der Metastruktur, das die verschiedenen Teile der Textbasis miteinander verknüpft (Abbildung 2).

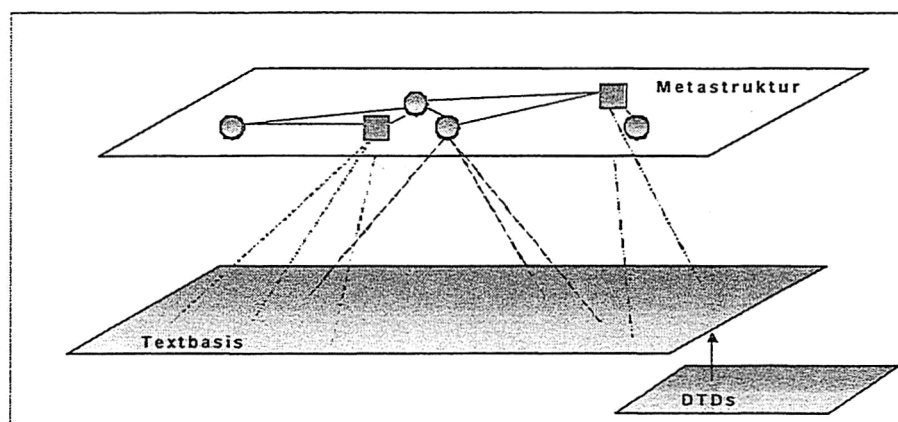


Abbildung 2: Bereiche des Informationspools

In seiner Rolle als zentrale Ressource für den Verlag dient der Informationspool nicht nur als Basis zur Herstellung der GKFA, sondern auch für verschiedene andere Thomas Mann Publikationen, wie beispielsweise eine Schulausgabe oder ein Themenband *Thomas Mann und Goethe*, die aus dem Datenbestand der Gesamtausgabe gewonnen werden können. Solche Abfragen für die Erstellung von Sonderpublikationen nutzen genauso das semantische Netz aus, wie es beispielsweise die elektronische Version tut, wenn sie es ermöglicht, nach den für Thomas Mann relevanten russischen Autoren des 19. Jahrhunderts zu fragen. Im Lektorat verschieben sich mit dieser Veränderung des Workflows traditionelle Aufgaben: Nicht mehr die spezifische Ausgabe steht im Mittelpunkt der Arbeit, sondern mehr und mehr die

Arbeit rund um den Informationspool. Er wird zur eigentlichen Textgrundlage und Arbeitsbasis, um die gesamte Textmenge zu edieren, miteinander zu verknüpfen und dabei auf den verschiedenen Ebenen konsistent zu halten. Der Informationspool ist nicht auf einen einzelnen Autor beschränkt. Er soll – ganz im Gegenteil – nach und nach um andere kommentierte und unkommentierte literarische Editionen sowie Einzelwerke erweitert werden.

Ein Kern – viele Gesichter

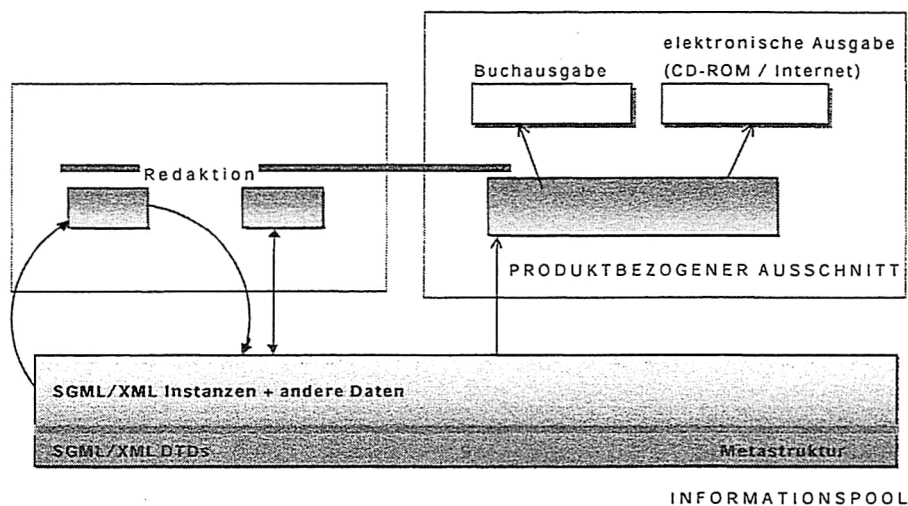


Abbildung 3: Schema der Publikationsarchitektur

Das Schema der Publikationsarchitektur (Abbildung 3) zeigt nun den Informationspool als zentrale Komponente des Gesamtkonzepts. Auf ihm setzen die Bereiche der redaktionellen Inhaltsbearbeitung und der Produkterstellung auf. Es werden entweder Inhalte neu nach den SGML/XML-DTDs strukturiert oder schon strukturiert vorliegende Dokumente bearbeitet. Hinzu kommt der Aufbau des semantischen Netzes als neue redaktionelle Aufgabe. Daneben werden für die Produkterstellung die Inhalte der jeweiligen Ausgabe bestimmt und aus dem Informationspool als produktbezogener Ausschnitt herausgezogen. Das kann beispielsweise ein einzelner kommentierter Thomas Mann Roman der GKFA sein, eine Sammlung von Thomas Mann und Heinrich Mann Briefen oder andere Ausgaben, die aus der Textbasis und der Metastruktur des Informationspools generiert werden. Aus einem solchen produktbezogenen Ausschnitt können dann Buchausgabe und elektronische Ausgabe mediengerecht entwickelt werden. Die inhaltliche Basis für die elektronische Präsentation kann dabei umfangreicher sein als die für das Buch. Die einheitliche gemeinsame Datenbasis ermöglicht jedoch als Resultat *eine Ausgabe in zwei medialen Ausprägungen*. Die Tiefe der Strukturierung nach semantischem Gehalt und die Ausdifferenziertheit der Metastruktur bestimm-

men dabei über die Möglichkeiten der mediengerechten Umsetzung, gerade für das elektronische Medium.

Das Archiv als Lösung

Um das Potential des neuen Mediums für die elektronischen Ausgaben der GKFA auszuschöpfen, lag es nahe, sich – neben der Entwicklung neuartiger Zugriffsmöglichkeiten auf die Inhalte – der multimedialen Möglichkeiten zu bedienen. Daran schloss sich jedoch unmittelbar die Frage an, welche multimedialen Erweiterungen im Rahmen einer wissenschaftlichen Werkausgabe angebracht sind. Der Phantasie waren da zunächst keine Grenzen gesetzt: verschiedene Porträtaufnahmen der Familie Mann, Filmaufnahmen zur Nobelpreisverleihung, die Skizze der Sitzordnung der Familie Buddenbrook beim Weihnachtsessen, Faksimiles der Bibel mit Thomas Manns Notizen zu seinem Josephs-Roman, Musikstücke, BBC-Radioansprachen im Originalton, eine Dokumentation der Schallplattensammlung Thomas Manns, Aufnahmen von Lesungen, Antwortbriefe seiner Briefpartner oder seine Notizbücher. Auf den ersten Blick schien sich eine Fülle an Möglichkeiten aufzutun – schöne neue Welt! War diese neue Welt wirklich so schön? Schon der zweite Blick trübte sie erheblich. Bei einer solchen Ausweitung im elektronischen Medium war es unklar, in welchem Verhältnis die elektronische Ausgabe zur Buchausgabe stehen sollte: Die Buchausgabe als wissenschaftliche Werkausgabe in strengem Sinne; die elektronische Ausgabe als *weniger* wissenschaftliche Werkausgabe, dafür aber mit höherem Fun-Faktor?

Der zentrale Anspruch der GKFA ist, die neue Referenzausgabe zu Thomas Mann zu sein. Referenzausgabe heißt, dass sie *die* verbindliche Ausgabe ist, um Thomas Mann Texte zu zitieren. Wenn nun aber die Konzeption der Buchausgabe von der der elektronischen abweicht, kann dann noch von einer Referenzausgabe in beiden Medien gesprochen werden? Aus dieser Fragestellung heraus wurde für die GKFA ein Konzept entwickelt, welches die Vorteile der möglichen Erweiterungen im elektronischen Medium mit der Geschlossenheit der inhaltlichen Konzeption in beiden Publikationsmedien der GKFA verbindet. Denn dass die GKFA in beiden Medien Referenzausgabe sein sollte, war eine zentrale Anforderung im Projekt.

Die Textbasis, die der Buch- und elektronischen Ausgabe zugrunde liegt, wird aufgeteilt in die GKFA und ein Archiv zur GKFA (Abbildung 4). Die GKFA besteht aus der Buchausgabe und möglichen, jedoch klar begrenzten Erweiterungen für die elektronische Ausgabe. Diese Erweiterungen könnten u.a. aus einigen wenigen Quellen bestehen, die aufgrund ihrer medialen Beschaf-

fenheit nicht im Buch publiziert werden können. Neben die GKFA wird ein Archiv gestellt, welches in Verantwortung des Verlages steht. In diesem Archiv finden sich Dokumente und Texte rund um das Werk Thomas Manns. Es nutzt die oben angesprochenen Möglichkeiten des elektronischen Mediums, nämlich Tondokumente zu präsentieren, schwer zugängliche Faksimiles einem breiten Publikum zur Verfügung zu stellen oder auch das Briefwerk Thomas Manns durch Antwortbriefe seiner Briefpartner anschaulicher zu machen.

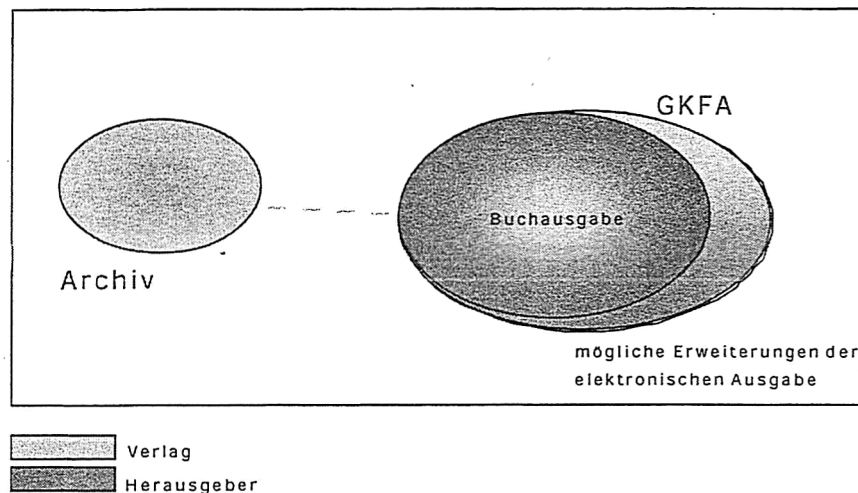


Abbildung 4: Bereiche der Textbasis

Das Archiv gibt dem Verlag die Möglichkeit, in eigener Verantwortung selten Gesehenes und Attraktives auf die elektronische Ausgabe zu stellen und damit dem Publikum einen Kaufanreiz zu bieten. Gerade dadurch, dass das Archiv ganz klar von der GKFA getrennt ist, ist es möglich geworden, eine Referenzausgabe in der oben beschriebenen Form zu haben und trotzdem multimediale Erweiterungen zu bieten.

Die Inhalte sind das Mass

Der größte Teil der Textbasis der GKFA ist SGML/XML strukturiert. Um die Daten für einen solchen Kontext aufzubereiten, muss das Modellierungskonzept der DTDs in erster Linie auf die Abbildung der inhaltlichen Einheiten und ihrer Zusammenhänge abzielen und nicht auf die hierarchische Strukturierung von Dokumenten. Die Inhalte sind damit in einer Struktur abgespeichert, die sie unabhängig von einer bestimmten medialen Präsentationsform macht. Die DTDs stehen in engem Zusammenhang mit der inhaltlichen Konzeption der Ausgabe. Sie stellen die Schnittstelle zwischen Konzeption und technischer Umsetzung dar, da sie die Anforderungen an die Ausgabe in einer formalen Sprache festhalten und damit die Basis für die technische Um-

setzung schaffen. Das DTD-Konzept selbst ist mit den bereits beschriebenen Teilen der Konzeption eng verwoben; auch hier wird die Textbasis von der Metastruktur und dem Archiv unterschieden.

Gerade der Anspruch einer inhaltlich ausgerichteten Strukturierung setzt allerdings voraus, dass die Entwicklung dieser Strukturen in enger Zusammenarbeit zwischen den inhaltlich Verantwortlichen, dem Lektorat, und den Informationsarchitekten erfolgt. Vergibt der Verlag diese Aufgabe an einen Softwarehersteller und zeichnet sich nicht selbst dafür verantwortlich, so kann die Modellierung zwangsläufig keine so hohe konzeptionelle Bedeutung einnehmen.

Neue Ausblicke mit Topic Maps

Ein wesentlicher Bestandteil des Thomas Mann Projekts ist der Aufbau eines semantischen Netzes für eine kommentierte literarische Werkausgabe. Dieses semantische Netz hat die Funktion einer Metastruktur und wird als Topic Map exportiert. Enzyklopädien und Wörterbücher sind geprägt von festen Strukturmustern und Formulierungsvorgaben. Ganz anders verhalten sich die Texte einer literarischen Edition; sie weisen gerade keine festen Strukturen auf, sondern leben von der Varianz der Sprache. Diese Lebendigkeit der Sprache erschwert eine automatisierte Inhaltserschließung, bzw. macht sie geradezu unmöglich. Um auf den Informationsgehalt der Texte dennoch mit technischen Mitteln gezielt zugreifen zu können, müssen die sprachlich ganz verschieden dargestellten Sachverhalte in die formale Notation eines semantischen Netzes bzw. einer Topic Map übertragen werden. Diese Übertragung ist sowohl ein intellektueller wie ein technischer Vorgang. Wird eine solche Formalisierung von Fakten geleistet ist, ist es möglich, komplexe Fragen zu recherchieren wie:

- In welchen Zeitungen wurden Rezensionen von Thomas Mann veröffentlicht?
- In welchen Essays äußert sich Thomas Mann zu anderen Schriftstellern seiner Zeit?
- In welchen Briefen schreibt Thomas Mann über seinen Zauberberg?
- Welche Autoren hat Thomas Mann am häufigsten zitiert?

Antworten auf diese Fragen finden sich in den Texten; diese sind jedoch ohne Aufbereitung technisch nicht zu erschließen. Wie man sich die Möglichkeiten, die sich aus dem Aufbau einer Topic Map ableiten, nun in der elek-

tronischen Ausgabe vorstellen kann, soll an nachfolgendem Benutzerszenario verdeutlicht werden.

Interessiert sich der Benutzer beispielsweise für den Briefwechsel Thomas Manns mit Schriftstellerkollegen zwischen 1889 und 1932, so kann ihm eine solche Gruppierung von Personen auf Grundlage der Modellierung des semantischen Netzes angeboten werden (Abbildung 5). Wählt er dann Gerhart Hauptmann aus, so werden alle an ihn gerichteten Thomas Mann Briefe gefunden.

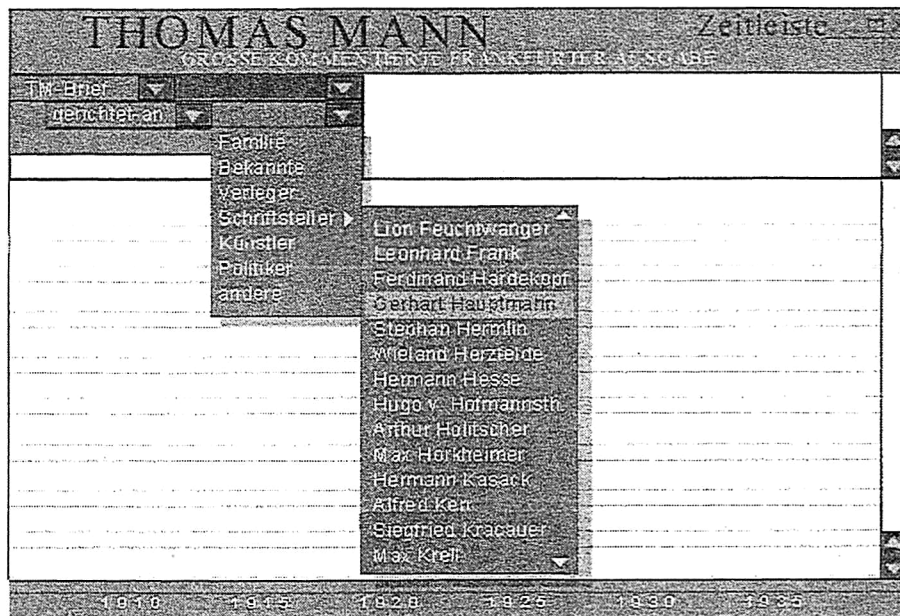


Abbildung 5: Suchanfrage

Das Suchergebnis kann wahlweise als einfache, chronologisch geordnete Liste angezeigt oder auch graphisch aufbereitet auf einer Zeitleiste präsentiert werden (Abbildung 6).

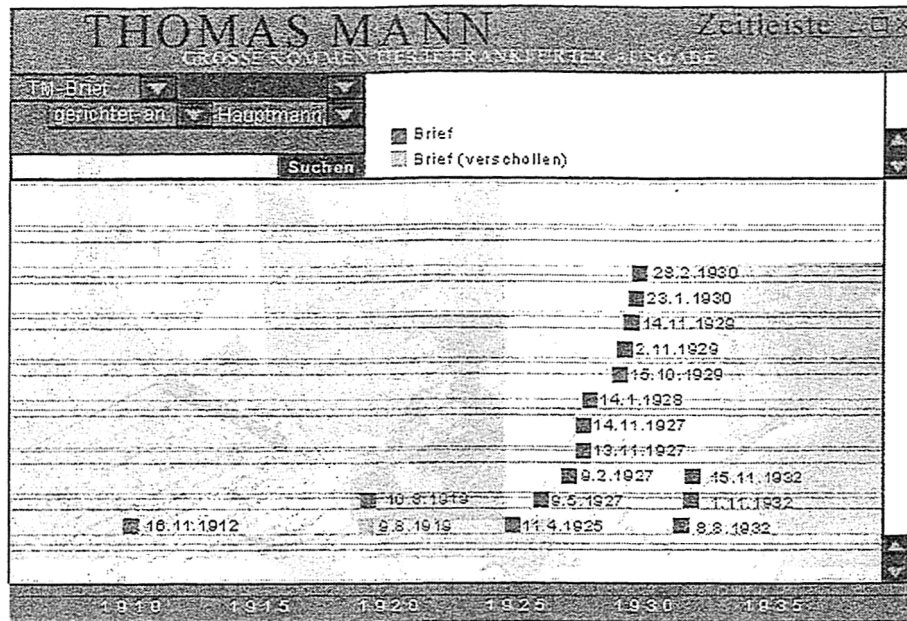


Abbildung 6: Ergebnis als Zeitleiste

Diese Zeitleiste wird dabei nicht vorab in der Redaktion erstellt und abgespeichert, sie wird vielmehr als Ergebnis der individuellen Anfrage automatisch generiert. Die Visualisierung auf der Zeitleiste zeigt auf den ersten Blick – sehr viel deutlicher als eine Liste das je könnte –, dass es genau einen Zeitraum intensiver Korrespondenz zwischen den beiden Schriftstellern gab. Möglicherweise steht der Anlass dieses regen Schriftverkehrs im Brief vom 11. April 1925, der den Beginn dieser Phase markiert. In diesem Brief entschuldigt sich Thomas Mann bei Gerhart Hauptmann dafür, dass er Charaktereigenschaften seines Kollegen in eine seiner Romanfiguren einfließen ließ. Der Brief sagt jedoch nichts darüber aus, um welche Figur es sich handelt. Auch eine solche Suchanfrage lässt sich auf Basis des semantischen Netzes formulieren. Ziel der konkreten Frage ist es, eine *Figur* zu finden, die in einem *Thomas Mann Werk* auftaucht und *Gerhart Hauptmann* zur *Vorlage* hat. Das Suchergebnis weist die gesuchte Figur als *Mynheer Peeperkorn* aus und ordnet sie dem Roman *Der Zauberberg* zu. Des Weiteren sind alle Kapitel und Unterkapitel des *Zauberbergs* aufgelistet, in denen *Peeperkorn* vorkommt. Abschließend werden die Materialien aufgelistet, die sich zu diesem Themenkreis im Archiv befinden (Abbildung 7).

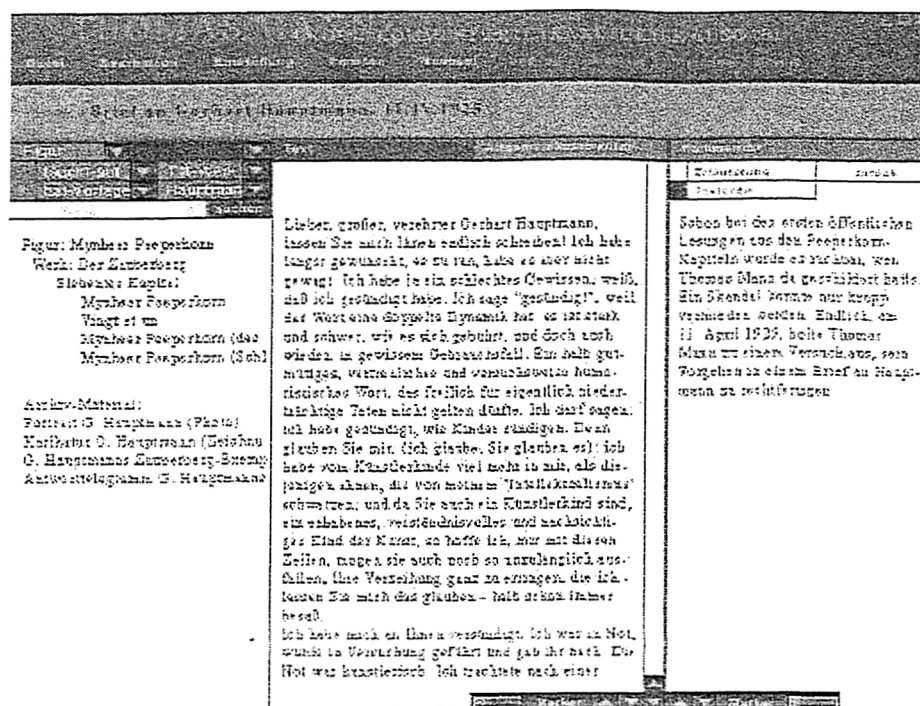


Abbildung 7: Brief mit neuer Suchanfrage und -ergebnis

Einblicke in den Aufbau

Um die eben beschriebenen Funktionalitäten möglich zu machen, muss ein formales Modell der Metastruktur, konzipiert werden, das in seiner inhaltlichen Ausgestaltung auf den Gegenstandsbereich der literarischen Edition maßgeschneidert ist. Auf der formalen Ebene teilt sich die Metastruktur in (1) die Konzeptebene und (2) die Ebene der Individuen. Auf der Konzeptebene findet sich das abstrakte Modell der Objekte und Relationen, in dem sowohl die möglichen Objekt- und Relationstypen als auch deren Verknüpfungsmöglichkeiten festgelegt werden. Diese Ebene des semantischen Netzes wird explizit aufgebaut und dient dann als Raster für die Einordnung der Individualobjekte. So wird beispielsweise auf der Konzeptebene angelegt, Personen nach ihren verschiedenen Rollen und Funktionen, als Familienmitglieder, Verleger, Schriftsteller, Politiker u.a. zu gruppieren, um dann das Individualobjekt Thomas Mann mit dem Konzept Schriftsteller verbinden zu können. Entsprechend wird mit dem Individualobjekt Zauberberg verfahren. In einem weiteren Schritt können diese beiden Individuen durch eine auf der Konzeptebene ermöglichte Relation miteinander verbunden werden, sodass dann das Faktum *Thomas Mann hat den Zauberberg geschrieben* im semantischen Netz abgebildet und damit auch recherchierbar ist (Abbildung 8).

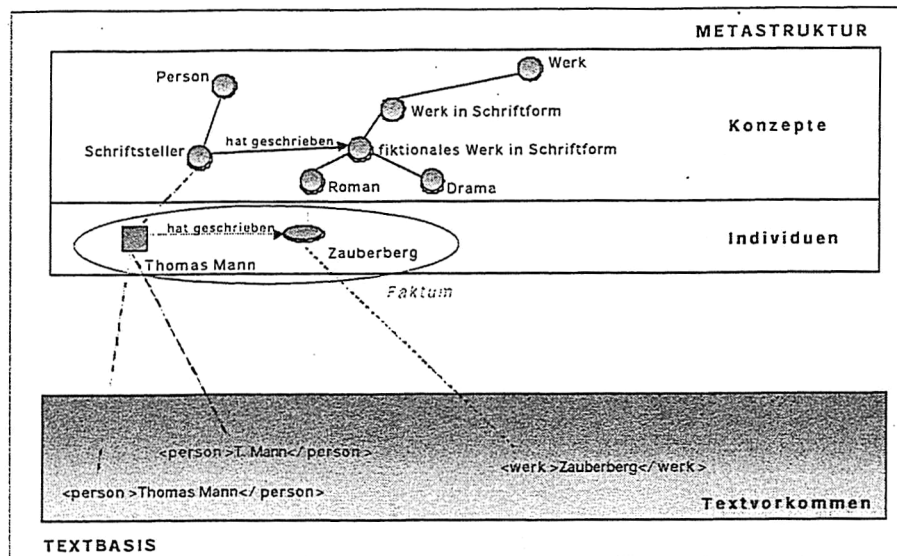


Abbildung 8: Schema des semantischen Netzes

Die Individuenebene ausschließlich manuell mit Individualobjekten aufzufüllen und diese zu verknüpfen, sodass Fakten gebildet werden können, wäre sehr zeitaufwendig und arbeitsintensiv. Als Alternative bietet sich an, Basisdaten automatisiert einzufüllen und diese redaktionell weiterzubearbeiten und zu ergänzen. Um zu einer solchen Basisfüllung zu kommen, lag es nahe, schon bestehende Register zum Werk Thomas Manns in das semantische Netz einzulesen.

Register als Basis

Register bieten sich deshalb an, weil sie innerhalb von Buchausgaben dem Leser die Möglichkeit bieten, nach bestimmten Begriffen die Texte zu erschließen. Sie sind außerdem stark formalisiert und daher auch gut automatisch auszuwerten. Für Thomas Mann liegen zahlreiche Register vor, die im Zusammenhang mit bereits früher publizierten Ausgaben entstanden sind. Diese Register werden für das automatisierte Füllen der Individuenebene des semantischen Netzes zunächst auf ihre Registerbegriffe reduziert, d.h. beispielsweise Seitenzahlen werden entfernt; in einem weiteren Konvertierungsschritt wird die Gestaltung der Register ausgewertet. Layout und Typographie setzen im Druck die Registerbegriffe nach einem strengen Regelsystem zueinander in Beziehung. So werden Fakten wie *Katia Mann hat Meine ungeschriebene Memoiren geschrieben* abgebildet. Dieses Faktum ist beispielsweise im Kurzke-Register [4] dadurch ausgedrückt, dass *Katia Mann* recte gesetzt ist, wie alle Personen, und *Meine ungeschriebenen Memoiren* kursiv und eingerückt darunter steht, wie alle anderen Werke auch. Durch solche festen Strukturen wird eine automatisierte und damit regelbasierte

Auswertung der Register möglich. Das Ergebnis dieser Auswertung ist ein semantisch ausgerichtetes XML-Dokument.

Diese XML-Auszeichnungen werden dann für die Erstellung von Teilregistern genutzt. Ein Teilregister gruppiert beispielsweise alle Personen mit ihren Verwandtschaftsbeziehungen (Abbildung 9). Durch diesen Filtermechanismus entstehen ganz spezifische Sichten auf das Gesamtregister, die eine Überarbeitung der Konvertierung vereinfachen, da Inkonsistenzen im Datenbestand leichter sichtbar werden. Im oben erwähnten Kurzke-Register sind beispielsweise *Personen* und *Institutionen* typographisch nicht zu unterscheiden. Da sie auch keine anderen Charakteristika aufweisen, die es ermöglichen, sie automatisch auseinanderhalten, können sie bei der Konvertierung nicht differenziert werden. Da in diesem Register wesentlich mehr Personen als Institutionen vorkommen, werden bei der automatischen Analyse alle Institutionen als Personen gefasst; diese falsche Auszeichnung muss anschließend manuell korrigiert werden (Abbildung 10). Nach einer solchen Überarbeitung werden die Register in das semantische Netz eingelesen; aus den semantischen Auszeichnungen des Registers werden dabei so viele Fakten wie möglich abgeleitet.

```
<?xml version="1.0"?>
<DOC>
<ENTRY>
<PERSON>Dohm, Ernst</PERSON><REL TYPE="Grossvater">Katia Mann</REL>
</ENTRY>
<ENTRY>
<PERSON>Dohm, Hedwig</PERSON><REL TYPE="Grossmutter">Katia Mann</REL>
</ENTRY>
<ENTRY>
<PERSON>Gide, Madame Charles</PERSON><REL TYPE="Mutter">Andr  ; Gide</REL>
</ENTRY>
<ENTRY>
<PERSON>Gide, Emmanu  le</PERSON><REL TYPE="Cousine">Andr  ;
Gides</REL>
<REL TYPE="Gattin">Andr  ; Gides</REL>
</ENTRY>
<ENTRY>
<PERSON>Helios</PERSON><REL TYPE="Sohn">Kleopatra VII.</REL>
<REL TYPE="Sohn">Marcus Antonius</REL>
</ENTRY>
<ENTRY>
<PERSON>Mann, Elisabeth, geb. Marthy</PERSON><REL TYPE="Grossmutter">Thomas
Mann</REL>
</ENTRY>
```

Abbildung 9: Teilregister Verwandtschaftsbeziehungen

```
<?xml version="1.0"?>
<DOC>
<ENTRY>
<PERSON>Baumgartner, Marie </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Bayerische Akademie der Sch&ouml;n&uuml;nste </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Bebel, August </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Beccaria, Cesare </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Becher, Johannes R. </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Bechstein, Ludwig </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Becker, Hansjakob </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Becker, Karl Heinrich </PERSON>
</ENTRY>
<ENTRY>
<PERSON>Beckett, Samuel</PERSON><WERK ALTNAME="Warten auf Godot">
<TITEL>En attendant Godot </TITEL></WERK>
</ENTRY>
</DOC>
```

Abbildung 10: Teilregister Personen und Institutionen

Alle in diesem Zusammenhang relevanten Fakten können jedoch nicht automatisch erkannt werden, z.B. kann der Schriftsteller nicht automatisch vom Komponisten, das literarische nicht vom musikalischen Werk unterschieden werden (Abbildung 11). Diese Fakten lassen nur mit Sachkenntnis in der Redaktion nacharbeiten.

Beteiligte

Es wurde vielfach deutlich, dass ein Vorhaben wie das hier dargestellte Projekt eine intensive Zusammenarbeit zwischen verschiedenen Beteiligten voraussetzt. Die konkrete Arbeitsgruppe im Thomas Mann Projekt setzt sich zusammen aus Mitarbeitern von Lektorat und Herstellung bei S. Fischer (u.a. Konzepterstellung, editorische Arbeit, Betreuung der Buchausgabe, Aufbau des semantischen Netzes, Datenarbeiten), VIA (u.a. Konzepterstellung, Aufbau der Informationsarchitektur, Kommunikationsschnittstelle zwischen Verlag und Softwarehersteller, Konzeption des semantischen Netzes, Datenarbeiten), Intelligent Views (u.a. Konzeption des semantischen Netzes, Entwicklung der redaktionellen Tools und von Werkzeugen rund um das semantische Netz, Entwicklung der elektronischen Ausgabe) und Clausen&Bosse (u.a. Buchausgabe, Datenarbeiten). Der Kern der Arbeits-

gruppe mit regelmäßigen Treffen bilden die Mitarbeiter von S. Fischer und von VIA.

Zum Schluss

Das Thomas Mann Projekt ist ein Beispiel dafür, wie ein zukunftsweisendes verlegerisches Konzept heute aussehen kann. Der S. Fischer Verlag hat die neue Thomas Mann Werkausgabe zum Anlass genommen, neue Strategien im Umgang mit Inhalten zu etablieren. Der Verlag stellt nach wie vor schöne Bücher her. Er hat sich aber auch auf den Weg gemacht, seine Inhalte so aufzubereiten, damit er für die neuen Medien gerüstet ist. Die neuen Publikationsmärkte wie CD-ROM, E-Book und Internet Publikation verändern sich beständig. Es ist eine „Forderung der Zeit“, diese Märkte mit qualitativ hochwertigen Produkten zu bedienen und auf ihre Veränderungen flexibel reagieren zu können.

Dieser neue Umgang mit den Inhalten geht einher mit Veränderungen im Verlag: neue Aufgaben entstehen, schon bestehende Aufgabenbereiche verzahnen sich stärker. Es ist eine weitere „Forderung der Zeit“, diese Veränderungen als Chance und Herausforderung zu begreifen.

[Dieser Artikel wurde in modifizierter Form in der iX 11/2000 veröffentlicht.]

Literatur

[1] Thomas Kamps; Christoph Obermeier; Klaus Reichenberger; Ingrid Schmidt, SGML für dynamische Publikationen – das Beispiel Fischer Weltalmanach, in: Wiebke Möhr; Ingrid Schmidt (Hg.), SGML und XML. Anwendungen und Perspektiven. Heidelberg 1999, 173-192.

[2] Eve Maler; Jeanne El Andaloussi, Developing SGML DTDs. From Text to Model to Markup, 1996.

[3] Katia Mann, Meine ungeschriebenen Memoiren, hrsg. v. Elisabeth Plessen und Michael Mann, Frankfurt a. M. 1976, 45f.

[4] Thomas Mann, Essays, hrsg. v. Hermann Kurzke und Stephan Stachorski, Frankfurt a.M. 1993ff.

[5] Hans Holger Rath, Mozart oder Kugel. Mit Topic Maps intelligente Informationsnetze aufbauen, iX 12/1999, 149-155.

[6] Ingrid Schmidt; Carolin Müller, Planning a new type of literary edition: the Thomas Mann Project, in: XML Europe 2000 (12-16 June 2000, Le Palais

I. Schmidt, C. Müller

de Congrès de Paris). Conference Proceedings, 83-97.
<www.gca.org/papers/xml europe2000/papers/s09-02.html>